8-2015

Population Structure of Late Blight (*Phytophthora infestans*) in Colombia and Ecuador and Downy Mildew (*Peronospora farinosa* f. sp. *spinaciae*) on Spinach in Arizona and California

Rebecca Marie Lyon
University of Tennessee - Knoxville, rlyon3@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Rebecca Marie Lyon entitled "Population Structure of Late Blight (*Phytophthora infestans*) in Colombia and Ecuador and Downy Mildew (*Peronospora farinosa* f. sp. *spinaciae*) on Spinach in Arizona and California." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Entomology and Plant Pathology.

Kurt H. Lamour, Major Professor

We have read this thesis and recommend its acceptance:

John K. Moulton, Rebecca T. Trout Fryxell

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Population Structure of Late Blight (*Phytophthora infestans*) in Colombia and Ecuador and Downy Mildew (*Peronospora farinosa* f. sp. *spinaciae*) on Spinach in Arizona and California

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Rebecca Marie Lyon
August 2015

Acknowledgements

I would like to thank all of those who contributed to the success of this thesis. First, I would like to thank my major professor Dr. Kurt Lamour and my committee members Drs. Kevin Moulton and Rebecca Trout Fryxell. I would also like to extend my gratitude to colleagues and contributors for the *Phytophthora infestans* project including, Dr. Sophien Kamoun, Dr. Kentaro Yoshida, Dr. Peter Kromann, Israel Navarrete, and Dr. Silvia Restrepo. I wish to acknowledge those who contributed to the *Peronospora farinosa* f. sp. *spinaciae* project including Drs. James Correll, Chunda Feng, Burt Bluhm, and Ainong Shi. Finally I would like to thank a fellow lab member Sandesh Shrestha.

Abstract

In this study, *Phytophthora infestans* and *Peronospora farinosa* f. sp. *spinaciae* populations were analyzed using single nucleotide polymorphisms (SNPs). In Ecuador and Colombia, *Phytophthora infestans* causes significant damage to potato and tomato and the epidemiology is known to be highly clonal. Our objective was to measure population structure within the context of this clonal epidemiology using both synonymous and nonsynonymous markers. Candidate SNP sites were selected by comparing the draft genomes of the Ecuadorian isolates EC1-3527 and EC1-3626. Genotypes were assessed directly from infected tissue using a targeted sequencing approach. A total of 54 polymorphic sites were assessed in 93 infected plant samples revealing 28 unique multi-locus genotypes. Our results were consistent with previous studies indicating two clonal lineages are dispersed across both countries and the epidemiology is driven by clonal reproduction. Interestingly, the SNP markers revealed potentially functionally important, asexually-derived, genetic variation in nonsynonymous loci – including asexual variation in genes encoding RXLR effector proteins. This highlights the potential for rapid evolution within the context of clonal populations. The second pathogen, *Peronospora farinosa* f. sp. *spinaciae* is an obligate pathogen that causes downy mildew on spinach and is considered the most economically important disease of spinach. The objective of the current research was to assess genetic population structure of known historical races and isolates collected in 2014 from production fields in Yuma, Arizona and Salinas Valley, California were assessed using candidate synonymous SNPs identified by comparing sequence data from reference isolates of known races of the pathogen collected in 2009 and 2010. Genotypes were evaluated using targeted sequencing on genomic DNA extracted directly from infected plant tissue. Genotyping 26 historical and 167 modern samples at 46 SNP loci revealed 82 unique multi-locus genotypes. The unique genotypes clustered into six groups. The majority of isolates collected in 2014 were genetically similar, regardless of source location. The historical samples representing several of races showed greater genetic differentiation. Overall, the SNP data suggests within field genotypic variation was primarily a product of asexual development in production fields, whereas genetic diversity in the historical collection may be influenced by sexual recombination.

Table of Contents

List of Tables

List of Figures

Introduction

Oomycetes are a class that belong to the kingdom Chromista (1). They have filamentous growth, acquire nutrition by absorption, and rarely have septa in hyphae resulting in coenocytic (multinucleate cells). Typically diploid in the vegetative state, oomycetes can reproduce sexually when gametangia (oogonia and atheridia) are produced and meet. Some oomycetes, including *Phytophthora infestans* and *Peronospora farinosa* f. sp. *spinaciae*, require two different mating types (A1 and A2 types) to sexually reproduce. Sexual reproduction results in the formation of thick walled spores called oospores which can survive extended periods outside of host tissue. Oospores produce hyphae which may immediately produce asexual sporangia. Sporangia are oblong spores that can be wind dispersed and germinate directly to serve as primary inoculum or, the presence of free water, can produce bi-flagelle swimming spores known as a zoospores (2).

Phytophthora species are often hemibiotrophic. The organism typically has a biotrophic phase at the beginning of the infection process, feeding on living host cells using specialized structures known as haustoria. At some point the infection switches to a necrotrophic phase where the host cells are killed and the nutrients are acquired directly (3). The naming of the genus *Phytophthora* followed extensive investigations following the infamous Irish Potato Famine (3). The famine, occurring in 1845-1846, was the result of the potato blight caused by the oomycete *Phytophthora infestans*. *P. infestans* has a broad host range in the Solanaceae family and currently causes billions of dollars in losses each year (4). While *P. infestans* is known to have played a role in the Irish Potato Famine, the origin of this organism is controversial. Two theories have emerged, one in which *P. infestans* is thought to have originated from the center of origin of potato in Peru and Ecuador. The other theory is that P. infestans in Mexico which is the current known center of diversity (5, 6).

Worldwide diversity of *Phytophthora infestans* varies by region. Until the early 1980s, Mexico was the only place where both the A1 and A2 mating types could be found, accounting for the high level of diversity indicative of sexual reproduction. The rest of North America has had low levels of diversity. A single A1 clonal lineage of late blight was detected in the United States and Canada in the early 1840s. This clonal lineage dominated until what is known as the “second migration” occurred in the 1970s. Subsequent migrations introduced the A2 mating type (4).

In Europe, there was a wide distribution of a single A1 clonal lineage thought to be first introduced in the late 19th century. The second migration event occurred in the early 1980s and introduced the A2 mating type to Europe. *P. infestans* may have been introduced to Africa in 1941 from the UK with secondary introductions possible in the late 1980s. While most isolates identified in Africa were of the A1 mating type, a single A2 mating type was isolated in Morocco. *P. infestans* in South America has shown low levels of diversity throughout the continent (4).

While *Phytophthora* species are typically hemibiotrophic, downy mildews are often biotrophic and feed on only living host cells. Downy mildew on spinach is caused by the organism *Peronospora farinosa* f. sp. *spinaciae* and is the single most economically important disease of spinach (7). *P. farinosa* is race-typed based on infection phenotypes of a standardized set of spinach cultivars (8). Using mitochondrial DNA, the population of worldwide isolates can be separated into two groups: Group I for Asia and Oceania and Group II that includes America, Europe, and Japan (9). In general, especially compared to the much-studied late blight of potato, there is very little known about the population structure of spinach downy mildew.

Chapter I
Population Structure of *Phytophthora infestans* in Colombia and Ecuador Described Using Single Nucleotide Polymorphisms

Introduction

Phytophthora infestans is a near-obligate hemibiotrophic oomycete that attacks a variety of Solanacae hosts (10, 11). It is heterothallic, requiring both the A1 and A2 mating types to complete the sexual stage and produce thick-walled, dormant oospores. Similar to other outcrossing Phytophthora species, individual strains of *P. infestans* carry a high load of heterozygosity (e.g. roughly every 100 to 200bp) and it is obvious when populations are impacted by meiosis as there are numerous multi-locus genotypes and the presence of both mating types. During wet and cool conditions, the asexual lifecycle is completed in a few days and devastating epidemics can develop quickly (10). Foliar symptoms include dark expanding lesions surrounded by water-soaked areas and hosts with woody tissues can have dark stem lesions (3). Each year, *P. infestans* results in billions of dollars in expenses and losses worldwide and the mechanisms by which it adapts to resistant hosts and/or chemical controls are poorly understood (11). Populations of *P. infestans* in Ecuador and Colombia have been analyzed using both phenotypic characters and genotypic markers. Phenotypes include allozymes, mating type and virulence on cultivated and wild *Solanum* species. Genetic studies have used restriction fragment length polymorphism (RFLP), amplified fragment polymorphism (AFLP), mitochondrial haplotype, and simple sequence repeat (SSR) markers (5, 10-16).

In the 1990s, a clonal lineage recovered from Ecuador had a unique allozyme and RFLP profile which differed from *P. infestans* lineages found elsewhere the world. The lineage, named EC-1, was the A1 mating type (15). In 2001, in Ecuador, two A2 isolates were collected on pear melon (*S. muricatum*) and RFLP and mitochondrial sequence analyses placed the isolates into a different lineage, dubbed EC-2. Furthermore, a study of wild and cultivated Solanum species in the late 1990s led to the designation of an additional type, EC-3 (17). While initially it appeared that *P. infestans* was host specific, in 2002, the US-1 lineage with an A1 mating type was recovered from pear melon in the same field as isolates with the A2 mating type; marking the first time both mating types were found on the same host in Ecuador (12, 18). Attempts to cross Ecuadorian lineages (US-1, EC-1, EC-2 and EC-3) *in vitro* proved difficult, suggesting a genetic barrier to sexual reproduction may exist among Ecuadorian lineages (19).

Isolates collected in Ecuador between 1995 and 2002, on a variety of cultivated and wild *Solanum* hosts, were found to be US-1, EC-1, or EC-3 using isozymes, mitochondrial haplotypes, RFLP, and/or AFLP fingerprinting. The AFLP analyses grouped isolates into two major clusters with the EC-1 and US-1 genotypes in one cluster and the EC-3 and isolates recovered from the *Anarrhichomenum* complex in the second cluster. The *Anarrhichomenum* complex is a group of wild Solanacae species that at the time had not been fully taxonomically characterized. Each host species appeared to be attacked by a primary lineage (e.g. the EC-1 lineage on *Solanum tuberosum*). Cultivated pear melon (*Solanum muricatum*) was the exception with both

the US-1 and EC-2 lineages (A1 and A2 mating types, respectively) both causing infection, sometimes at the same location, as described above (5).

On *S. ochranthum* in Ecuador, isolates of *P. infestans* collected in 1997-2004 were all the A1 mating type. Three groups were identified based on mtDNA haplotype, isozyme profile and RFLP and SSR genotypes. Group A was not associated with any known lineage and appeared specific to *S. ochranthum*. Group B is a subgroup of the US-1 genotype and Group C was identical to EC-1 (13). A report from Colombia in 2006 based on analyses of mitochondrial haplotypes and SSR markers in isolates from a variety of hosts indicates most were the A1 mating type and the EC-1 lineage and there were few polymorphisms detected within the EC-1 lineage. On *Physalis peruviana* (Cape gooseberry), one isolate was A2 and identified as the US-8 clonal lineage. On potato, CO-1, CO-2, and EC-1 lineages were detected and on Cape gooseberry, CO-1 and US-8 were identified (14).

P. infestans isolates collected between 1999 and 2006 in Columbia and Venezuela had low levels of genetic variation in four nuclear and one mitochondrial gene and there was limited gene flow among the countries (11). In addition, low levels of variation have been reported for the EC-1 lineage in Ecuador using SSRs (15, 16). In Colombia, no variation was seen on *P. infestans* isolates on potato hosts when 19 single nucleotide polymorphisms (SNPs) were analyzed in regions flanking SSR markers (20).

Most population studies of Phytophthora have been conducted on isolates recovered from infected plant material followed by growth on nutrient media to recover single zoospore or hyphal tip-derived isolates. Although this approach was traditionally thought to have advantages (e.g. an isolate in culture is thought to be a homogenous and genetically “pure” individual) and for many “anonymous” markers systems (e.g. AFLP and RAPD) this approach is required; it is becoming increasingly clear that isolates can change rapidly, and in some cases, quite dramatically, while growing on nutrient media (21). These changes include changes in mating type, changes in virulence, loss of pathogenicity and at the genetic marker level, often manifest as loss of heterozygosity (LOH) – a phenomenon where previously heterozygous markers spontaneously switch to homozygosity (22). Our goal was to avoid genetic changes common to *in vitro* growth and to investigate the frequency of asexually-derived genetic changes occurring within the context of a known clonal epidemiology. To accomplish this, genotypes are assessed using genomic DNA extracted from discrete lesions and the markers used in the study are polymorphic between two previously described isolates of *P. infestans* from Ecuador. In addition the markers used here fall within genes and include both synonymous and nonsynonymous mutations (roughly 50% of each type).

Materials and Methods

Sample collection and DNA extraction

Infected plant material, exhibiting signs and symptoms typical for infection by *P. infestans*, was collected in March 2013. Each sample in our study is derived from a single infected leaf on a single plant. The greenish part of the lesions were excised and placed directly into RNA Later solution (Life technologies) in the field. Hosts include *S.*

betaceum (tamarillo), *S. caripense* (tzimbalo), and multiple varieties of *S. tuberosum* (potato). Samples were collected from the Ecuadorian provinces of Carchi, Chimborazo, Tungurachua, Pichincha, and Napo. Colombian samples were collected around Bogota in seven potato fields. DNA was extracted directly from infected tissue using the Omniprep kit (BG Biosciences) according to the manufacturers instructions.

Candidate SNP Identification and assessment in field samples

Candidate single nucleotide polymorphisms (SNPs) were identified based on re-sequence data from two isolates recovered from Ecuador (isolates EC1_3527 and EC1_3626) according to the method of Yoshida et al. 2013 (23). EC1_3527 and EC1_3626 were recovered in 2002 and 2003 from *Solanum andreanum* and *S. tuberosum*, respectively. Both isolates have been maintained in long term storage in liquid nitrogen. A total of 96 candidate SNPs in 95 genes were selected for amplification and use for population structure analyses 34 were predicted to be synonymous and 62 were nonsynonymous. Of the 96 SNPs, 88 are heterozygous SNPs in one or both of the EC1 isolates. Six homozygous SNPs were different from the T30-4 isolate, the first *P. infestans* isolate to have its genome fully sequenced and annotated (22). Two SNPs had a fixed (homozygous) difference between the two EC1 isolates (22).

For targeted sequencing, individual primer pairs were designed using Batch Primer 3 (http://probes.pw.usda.gov/batchprimer3/) at default settings for generic primers with total amplicon size set as an optimum of 100bp with the amplified region containing the target SNP (Table 4, appendix). The primer sequences and genomic DNA were submitted to Floodlight Genomics (FG, Knoxville, TN) for processing using a Hi-Plex targeted sequencing approach as part of FG's Educational and Research Outreach Program (FG-EROP) which provides services at cost (no profit) for research projects with no commercial components to the work (24). The Hi-Plex approach pools primers to PCR amplify targets and adds a barcode sequence and a sequencing adapter during the amplification process. The resulting target library is then sequenced on a next generation device (e.g. Ion Proton or Illumina device). In this study, the libraries were sequenced on an Ion Proton device.

Binned raw sequence data obtained from FG was processed using CLC Genomics Workbench 7.5 (www.clcbio.com, Qiagen, Arhus) to map the sample-specific sequences to a reduced representation reference genome containing only the target sequences. SNP genotypes were assigned to target sites with at least 20X sequence coverage and sites with <15% alternate allele frequency were considered homozygous genotypes.

Population structure was examined using a Bayesian Markov chain Monte Carlo (MCMC) clustering model using the program STRUCTURE v2.3.4 (25, 26). The admixture model using sampling locations as the prior was used with the lambda fixed at 1 with a burn-in of $2x10^5$ and $2x10^5$ iterations. There were 20 independent runs for each K value (number of clusters) from 1 to 10. The value of K that best fit our data was selected using the ΔK statistic calculated by Harvester (27). Arelquin was used to

calculate population pairwise F_{ST} within each structure hypothesized cluster and principle coordinates analysis of the clusters was performed using GenAlEx (28-30).

Phylogenetic analyses were conducted with MEGA 6 using the Maximum Likelihood method based on the General Time Reversible model (31). The bootstrap consensus tree was inferred from 500 replicates. Evolutionary rate differences among sites were modeled with a discrete Gamma distribution with the rate variation model allowing for some sites to be evolutionarily invariable (G + I). Initial trees for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach. Three additional genotypes were included in all analyses; the EC1-3527 and EC1-3626 isolates recovered from Ecuador and used for SNP selection and the T30-4 isolate which was used to make the reference genome of *P. infestans* (22, 23).

Results

Targeted sequencing and genotypes

The DNA from 113 infected plant samples were submitted to FG for targeted sequencing. Twenty of the samples amplified poorly and were not included in the study. The failure could be due to low *P. infestans* DNA content (e.g. most of the DNA in the sample was from the plant) or degradation following the sample storage and shipping and subsequent extraction. Approximately 10 million reads were generated by FG for the remaining 93 infected plant samples. These included 34 samples from Ecuador and 59 from Colombia (Table 1). Not all of the original 96 targets amplified with a minimum of 20x coverage to be included in the study. This is not uncommon as the primers were not calibrated or optimized and further work will be needed to either develop a separate assay for the low-performing primer combos or additional optimizations will be needed. There were 69 SNP markers with no missing data for all 93 samples for a total of 6417 genotypes. The average sequence coverage for each genetic target was >100X, ranging from 20 to 4000X. As outlined in the Hi-Plex approach, this range can be optimized based on titrating the primer concentrations. For our purposes, the variation in sequence coverage was compensated by over-sequencing. For the 69 successful markers, 15 were homozygous in all samples and the remaining 54 polymorphic markers were used for all further analysis (Table 5, appendix).

Unique genotypes and clonal lineages

A total of 28 unique multi-locus genotypes were identified. Of these, 10 identical genotypes were found more than once and designated G1-G10; ranging in frequency from 3 to 19 instances (Table 2). Genotype G1, recovered from 20% of the samples, was found on both *S. phureja* and *S. tuberosum*, and also recovered at the lowest and highest altitudes (2300 and 3400 meters, respectively) included in the study. The 28 unique multi-locus genotypes plus the reference genotypes from EC1-3527, EC1-3626 and T30-4 were used for the following population structure and phylogenetic analyses.

Table 1. Summary sample data.

Region, Country	Hosts	# of Samples	Approx. Altitude (m)
Pichincha, Ecuador	*S. betaceum, S. tuberosum*	21	2300-3400
Carchi, Ecuador	*S. tuberosum*	5	2300-3000
Chimborazo, Ecuador	*S. tuberosum*	3	2600-2700
Tungurachua, Ecuador	*S. betaceum, S. caripense, S. tuberosum*	3	2100
Napo, Ecuador	*S. tuberosum*	2	3200
Bogota, Colombia	*S. tuberosum*	59	2600

Table 2. Summary data for identical multi-locus SNP genotypes recovered more than once.

Genotype Name	# of samples	Hosts	Region	Approx. Altitude (m)
G1	19	*S. phureja, S. tuberosum*	Pichincha, Chimborazo, Carchi, Napo, Bogota	2300-3400
G2	16	*S. phureja, S. tuberosum*	Bogota	2600
G3	9	*S. phureja, S. tuberosum*	Bogota	2600
G4	7	*S. tuberosum*	Pichincha	3000
G5	7	*S. tuberosum*	Pichincha, Bogota	2600-3000
G6	4	*S. tuberosum*	Bogota	2600
G7	4	*S. tuberosum*	Bogota	2600
G8	4	*S. tuberosum*	Carchi, Bogota	2600-2700
G9	3	*S. tuberosum*	Pichincha, Bogota	2400-2700
G10	3	*S. tuberosum*	Carchi, Bogota	2600-2800

Structure analysis grouped the genotypes into three clusters with ΔK = 3 (Figure 1). T30-4, the *P. infestans* isolate used to produce a reference genome fell into the third cluster and EC1_3626 and EC1_3527 fell into clusters one and two respectively (Figure 2. All Colombia samples clustered into cluster one and Ecuadoran samples were present in all three clusters. Samples on the host *S. phureja* are found in cluster one.

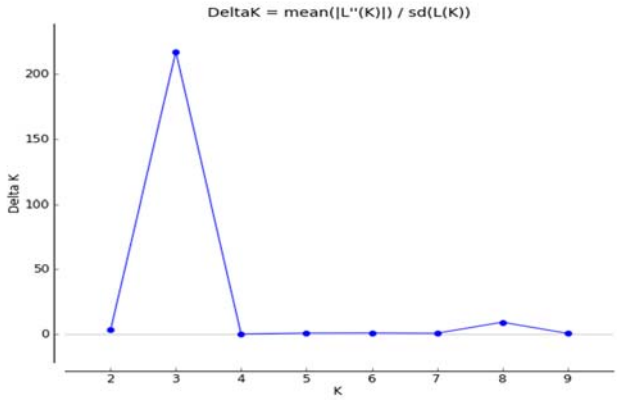


Figure 1. Graph showing DeltaK for *P. infestans*

A maximum likelihood tree was created using MEGA6 (Figure 3). The samples are colored based on the cluster identified by STRUCTURE; cluster 1 is red, cluster 2 is green, and cluster 3 is blue. Each cluster is a monophyletic group, forming three separate clades. A principle coordinate's analysis (PCA) confirmed the three clusters (Figure 4). Doing a population pairwise F_{ST}, the distance between cluster 1 and 2 is 0.157, cluster 1 and 3 is 0.357, and cluster 2 and 3 is 0.375, at a significance level of 0.05, showing significant differences between the three clusters.

Functionally important SNPs

Of the 54 polymorphic markers used in the study, 30 are predicted to be nonsynonymous and 24 are predicted to be synonymous (non-amino acid changing). Almost 40% of the markers (21 of 54 markers) are for SNP sites that fall into genes predicted to encode RXLR-type secreted proteins. Of these, 15 are predicted to be nonsynonymous and contain amino-acid changing mutations. Although little is known concerning the function of these proteins (or these exact amino acids) during the infection of potato in Ecuador and Columbia, it is interesting to note that differences do exist between the two main clonal groups active in those countries. For example, markers 1.2_3934005 and 1.5_1367534 are fixed for homozygosity in the group containing EC1-3626 and fixed for heterozygosity in the group containing EC1-3527. The other nonsynonymous sites are either fixed similarly (hetero or homo) for both groups, or have a mixture of homo and heterozygous genotypes (e.g. markers 1.11_1535743, 1.14_2071843, 1.95_137980 and 1.21_1258416)

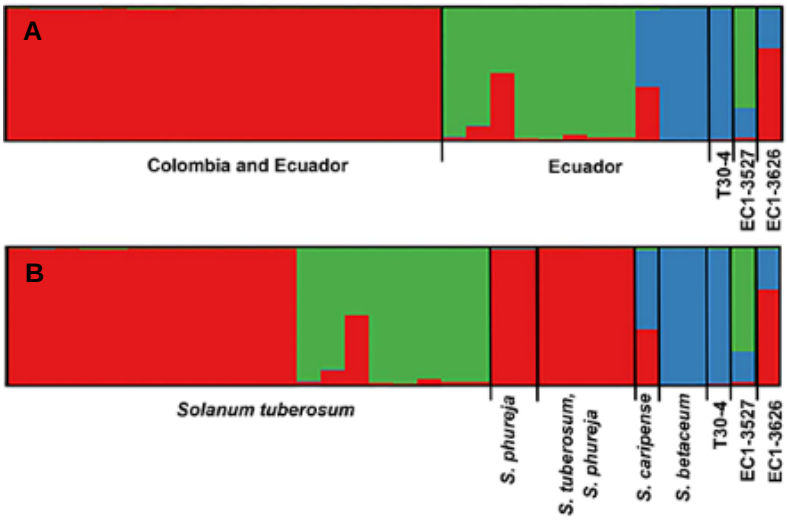


Figure 2. STRUCUTRE plot for *P. infestans*

STRUCTURE plots of assignment probabilities of 31 unique genotypes. Each genotype is represented by a bar indicating likelihood of membership in cluster 1 (red), cluster 2 (green), and cluster 3 (blue). Populations are defined by A) location or B) host. Samples from *S. tuberosum* are in both cluster 1 and 2. All samples collected on *S. caripense* and *S. betaceum* are in third cluster.

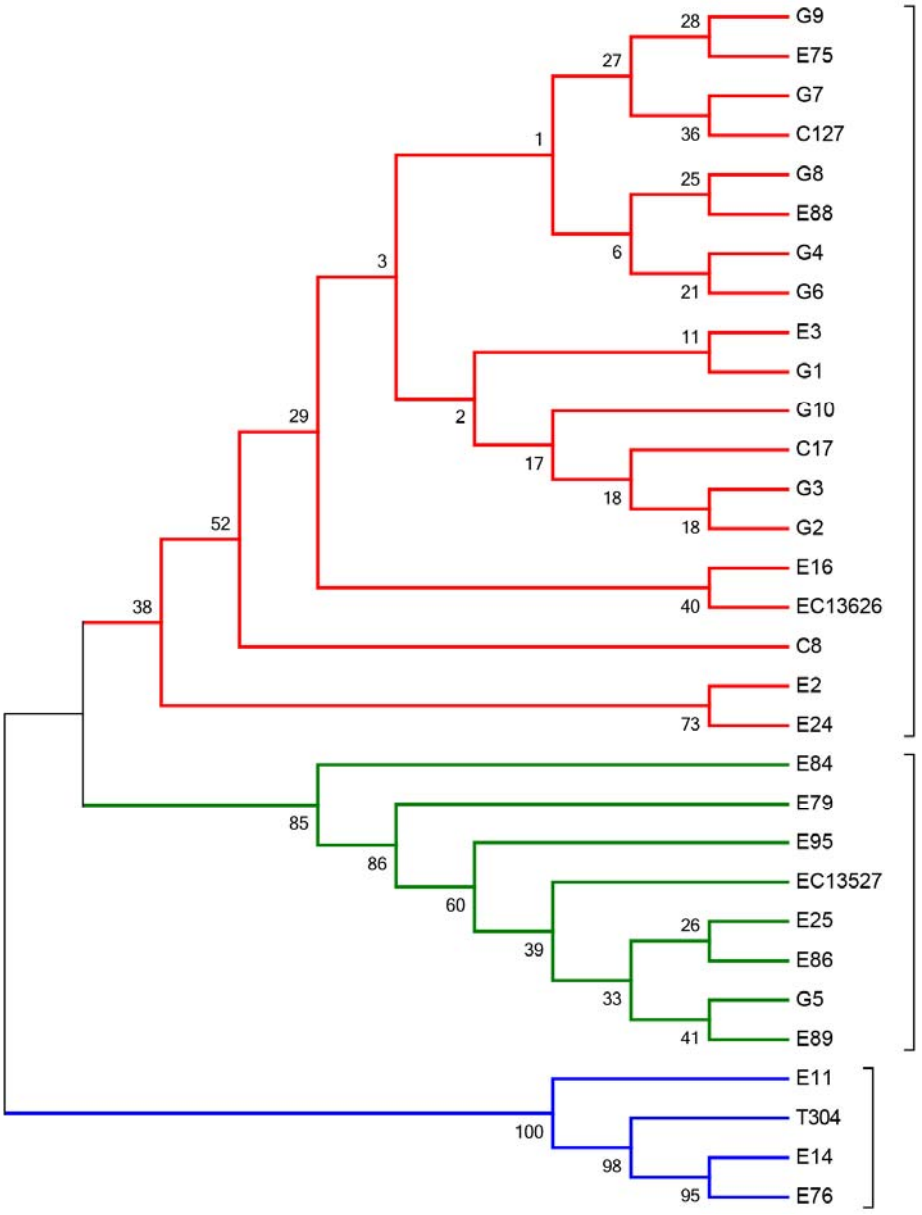


Figure 3. Maximum Likelihood tree created by MEGA6 for *P. infestans*
Each clade is colored based on STRUCTURE cluster. Cluster 1 (red), cluster 2 (green), and cluster 3 (blue).

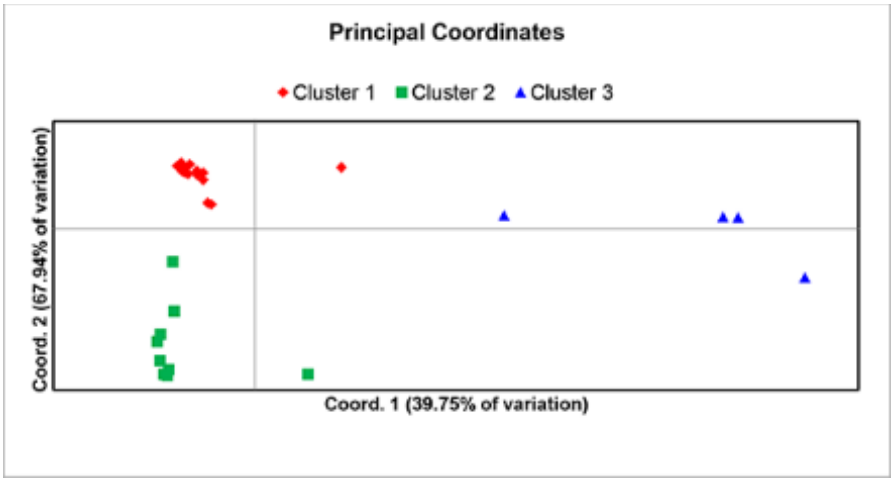


Figure 4. Principle Coordinates Analysis (PCoA) for *P. infestans*

The PCoA analysis confirmed three genetic clusters corresponding to the differences found using Structure and phylogenetic analyses. Each color represents a STRUCTURE cluster. Cluster 1 (red), cluster 2 (green), and cluster 3 (blue).

Allele frequencies and ploidy

The allele frequencies for heterozygous SNPs varied more than expected for a typical diploid organism where the alternate alleles are expected to form a distribution around 50%. For many of the heterozygous markers, the allele frequency distributions were found closer to 33% and 66%. There were some exceptions where the allele frequencies were distributed around 50%, but this was the exception and not the rule. Previous reports on *P. infestans* indicate some populations are triploid and it appears that may be the case for the two clonal groups causing damage on potato in Ecuador and Columbia (23, 32, 33). For example, the samples in cluster 1 (which had sufficient samples to more easily chart trends) are all heterozygous for markers 1.10_120399 and 1.72_552467 (Figure 5). Marker 1.10_120399 shows an allele frequency of around 50%. However, marker 1.72_552467 has allele frequencies distributed around 33% for one allele and 66% for the other allele. It is not clear, based on this data, if all the isolates are triploid at all locations in their genomes and further studies are warranted.

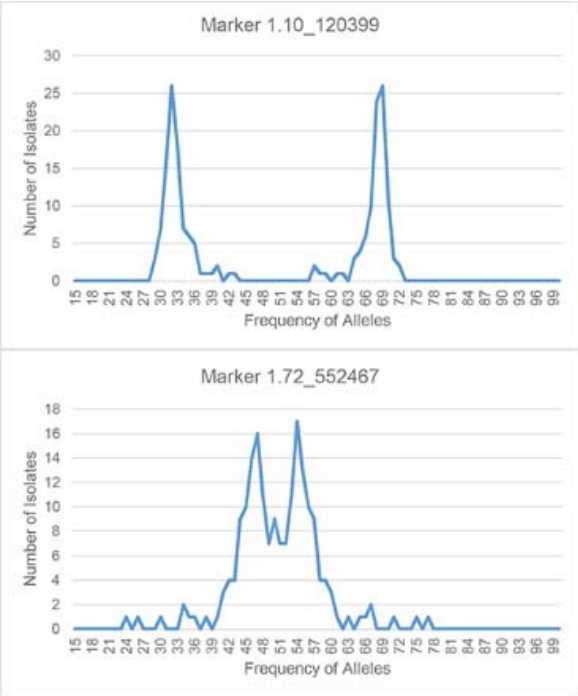


Figure 5. Histograms showing the frequency of alleles for two markers.

Marker 1.10_120399 shows a bimodal distribution at around 33% and 66% while marker 1.72_552467 is an example of a more unimodal distribution around 50%.

Discussion

Here we employ a panel of 54 anonymous and synonymous SNP markers to investigate the population structure of *P. infestans* in Ecuador and Colombia. Previous studies, using a variety of markers, including a limited number of SNPs, indicate the field populations are highly clonal with a limited amount of sub-clonal diversity Our goal was

to investigate the extent of sub-clonal diversity within the context of field epidemics using a larger battery of markers. Overall, our results support the previous studies indicating two clonal lineages are responsible for disease on a variety of potatoes in these countries. Although we only have a small number of isolates from non-potato hosts, our data indicates there is a third clonal lineage which attacks non-potato hosts. Our study differs from previous work by the inclusion of synonymous and nonsynonymous markers of which a relatively large proportion (40%) are considered RXLR class of secreted proteins, known as disease effectors (34). The so-called RXLR proteins are known to be important components for virulence and have been directly linked to interactions with R-genes in potato. The finding that nonsynonymous polymorphic sites can vary within the context of a clonal epidemic is interesting and potentially useful knowledge.

Also consistent with previous studies, was the presence of a third clonal lineage, distinct from the two causing disease on potato. The third group included all samples of *S. caripense* and *S. betaceum* and the reference genome of T30-4. The inclusion of the reference genome in this group may be artificial as the markers in this study were intentionally chosen to differentiate the Ecuadorian lineages and many of the markers are homozygous. The third group did not include any samples collected from potato plants or the EC-1 references.

P. infestans varies in regard to ploidy with isolates being diploid, triploid and trisomic (32, 35). It appears the diploid condition is not the standard in the populations described here and how this impacts or possibly even drives the observed mitotically-derived variation will require additional studies.

And finally, the multi-locus genotypes of field samples were directly evaluated using DNA extracted from lesions. By avoiding genetic changes known to occur during *in vitro* growth, this allowed a direct assessment of the dynamics of asexually derived genetic change and overall evolution in the field populations. Unlike RNA sequencing, the targeted sequencing approach used here is based on genome DNA and is not influenced by level of gene expression. Further studies on a finer scale, both within the genome and in field populations, will be valuable to better understand the mechanisms underlying the extensive, and potentially important, changes outlined here.

Chapter II
Downy Mildew in Arizona and California

Introduction

Spinach (*Spinacia oleracea*) is believed to be native to the Middle East and has been cultivated for more than 1,300 years (36, 37). It became established in Europe in the 11th century and was introduced to the new world and southern hemisphere in the 16th century (38).

Global demand for spinach has increased substantially during recent decades. In 1961, world production of spinach was 2.96 million tons. By 2012, production increased seven fold to 21.66 million tons worldwide. In the United States, spinach production increased from 200k tons in 1961 to 354k tons in 2012 (39). Per capita consumption of fresh market spinach increased from 0.3 kg/person in 1995 to 1.0 kg/person by 2005 (8). The increased demand largely due to the popularity of pre-packaged baby spinach leaves and the health benefits associated with spinach including vitamins C and A, lutein, iron, folic acid, and magnesium (40, 41).

Increased demand for spinach has dramatically influenced production practices. In California, planted acreage has expanded rapidly during the last two decades, 12-month production cycles are now common in some regions of the state, accompanied with high-density plantings, and reduced use of rotation crops. Furthermore, spinach seed usage has changed from the use of approximately 40,000 seed per acre to 40 million seeds per acre, dramatically changing plant population densities. As a result, the higher plant population densities have influenced disease pressure. Increased market demand has also affected the spinach seed production industry. Although some open pollinated lines of spinach are produced, hybrid seed represents about 90% production and is derived from inbred lines with different disease resistance characteristics (7, 42).

Peronospora farinosa f. sp. *spinaciae* causes downy mildew disease in spinach, one of the most economically important diseases constraining spinach production (7). It is an obligate, heterothallic pathogen with two mating types (P1 and P2) (43). Initial symptoms of downy mildew on spinach include yellow, irregular and chlorotic lesions on leaves. Blue-gray sporangia form when conditions are wet or during periods of high humidity (37). If conditions are dry, lesions turn white or tan (44). Sporangia are dispersed by wind and rain splash, and cause secondary cycles of disease development (37).

Isolates of *P. farinosa* f. sp. *spinaciae* can be classified into distinct races based on infection phenotypes on a standardized set of spinach cultivars (45). Rate of classification of new races of *P. farinosa* f. sp. *spinaciae* has exponentially increased with 15 races described (46). The first race of *P. farinosa* f. sp. *spinaciae* was identified in 1824, and in 1950, resistance to race 1 was determined to be conveyed by a single dominant gene. Race 2 was described in 1958 and soon thereafter resistance to both race 1 and race 2 was discovered (7). While resistance to race 1 and 2 was initially thought to be controlled by a single gene, resistance against races 1 and 2 was later demonstrated to be controlled by two closely linked genes (47). In 1976, race 3 was

identified and resistance to it was incorporated into commercial hybrids two years later (7). Race 4 was identified in 1990 in the United States (7, 48) and resistant hybrids were introduced soon after (36). Since 1990, new races of *P. farinosa* f. sp. *spinaciae* have been identified at an exponential rate (8, 45, 46). In 2000, a standardized set of race differentials for spinach was established by the International Seed Federation (ISF) to address confusion regarding races 5 and 6 in United States and Europe (49). Currently, the International Working Group on Peronospora (IWGP) meets annually to evaluate emergence of novel races and their potential importance to global spinach production. Multiple laboratories run parallel tests and, if the results constantly indicate unique virulence, a new race number is assigned (46).The exact cause for the rapid emergence of new races is unknown, but is postulated to result from intense selective pressure associated with new spinach production practices (7).

Management of spinach downy mildew is problematic, as chemical control options are limited. Cultural controls include crop rotation, reduction of “green bridges”, and efforts to reduce free moisture on leaves. The epidemiology of seedborne inoculum from seed remains to be fully evaluated. Oospores have been detected in washings of seeds and infected seedlings were produced from seed (43). If oospores are plowed under in field they are able to survive in the soil for up to one year. Oospores associated with infected seeds can potentially survive for two years (42).

The relationship between detection levels and the role of seedborne inoculum in field epidemics is currently unknown (40). A multiplex real-time PCR assay was recently created to detect various pathogens associated with spinach seeds, including *P. farinosa* f. sp. *spinaciae*. When used to assess seed lots, a high percentage of seeds tested positive for contamination via PCR, but infected seedlings were not observed in grow-out studies. This discrepancy may reflect the presence of inviable propagules of the pathogen, lack of seed transmission among certain isolates, or inadequate conditions for seed transmission in the experimental conditions (40).

To better understand the emergence of new races of *P. farinosa* f. sp. *spinaciae*, assessments of population structure on various spatial and temporal scales are critical. For example, new races could originate via spontaneous mutation among prevalent genotypes, or could originate from selective sweeps due to the introduction of new genotypes on infected plant material. Initial population genetics studies in *P. farinosa* f. sp. *spinaciae* were based on assessments of rDNA sequences and suggested low levels of genetic diversity among isolates (38, 40). However, subsequent analyses based on mitochondrial DNA allowed isolates to be classified by geographical origin (Group I for Asian and Oceania and Group II for American, European, and two Japanese samples). All base substitutions in the two mitochondrial genes tested were transitions, suggesting a recent genetic divergence. Because *P. farinosa* f. sp. *spinaciae* is an obligate pathogen, the disease is postulated to have spread with the crop and is thought to have originated in a small geographic area (9). However, the limited amount of genomic resources for *P. farinosa* f. sp. *spinaciae* has hindered efforts to utilize more sophisticated molecular markers to assess genetic diversity in the context of evolutionary biology.

Our goal was to develop novel single nucleotide polymorphism (SNP) markers for *Peronospora farinosa* f. sp. *spinaciae* and to characterize the population structure in field populations in Arizona and California and in a historical panel of race types.

Materials and Methods

Sample collection and genomic DNA extraction

Infected leaves were collected from spinach production sites in Yuma, Arizona in March of 2014 and from the Salinas Valley in May of 2014 and stored for 3 to 7 days on ice or in a refrigerator at 4C. A disc (approximately 7mm) of infected tissue was excised from each leaf and processed for genomic DNA extraction in a 96-well plate as previously described (50). Briefly, this involves freeze drying the tissue in a 2ml deep-well block followed by disruption into a powder and subsequent genomic DNA extraction using a silica based approach.

Historical race-type isolates, maintained at the University of Arkansas, were included in the study. The races have been previously characterized and maintained as references for phenotypic assays. Sporangia are routinely produced on susceptible plants grown in a growth chamber inoculation assay and washed off and freeze dried prior to DNA extraction as outlined above.

Candidate SNP identification

Genomic DNA was extracted from the *P. farinosa* f. sp. *spinaciae* isolates, UA2209 and UA4410 (race 12 and 14 respectively) and sequenced using 100bp paired-end Illumina sequencing on a HiSeq 2500 device (BGI, China). The sequencing is part of a larger effort to develop a reference genome for *P. farinosa* (underway and to be reported separately). Here the sequence data was exclusively used to identify putative SNP sites. Raw sequence data (≈145M paired reads per isolate) was assembled *de novo* into contigs using CLC Genomics Workbench 7.5 (www.clcbio.com, Qiagen, Arhus) at default settings. All further sequence manipulation, mapping and genotyping were accomplished with CLC. The resulting 300 largest contigs for isolate 2209 were processed further to identify all open reading frames (ORFs) greater than 1000 amino acids (3000bp). Of these, the largest 100 contigs with ORFs showing high predicted protein similarity, as assessed using Blastx in CLC, to other oomycete organisms (e.g. *P. infestans*) were used as a reference genome. The raw data from isolates 4410 and 2209 was then mapped separately to the 100 contig reference genome using settings of 90% similarity and 90% sequence coverage to identify heterozygous positions. Heterozygous positions were assigned when the alternate allele frequency was >15% and <85%. Polymorphic SNP sites that fell within the 1000 amino acid putative genes and are predicted to be synonymous (non-amino acid changing) were then catalogued and used to choose 72 SNPs residing on discrete contigs for further analyses.

Targeted sequencing and genotype assignment

Individual primer pairs were designed using Batch Primer 3 (http://probes.pw.usda.gov/batchprimer3/) at default settings for generic primers with total amplicon size set as an optimum of 100bp with the amplified region containing the

target SNP (Table 6, appendix). The primer sequences and genomic DNA were submitted to Floodlight Genomics (FG, Knoxville, TN) for targeted sequencing. Floodlight Genomics uses a Hi-Plex approach to amplify targets (80-100bp) in a multiplexed PCR reaction and then generates sample specific sequences using a next-generation sequencing device (e.g. Ion Proton or Illumina MiSeq). The targeted sequencing was done at no charge as part of the FG Educational and Research Outreach Program and was obtained using an Ion Proton (Life Technologies). Binned raw sequence data obtained from FG was then used to map the sample-specific sequences to a reduced representation reference genome containing only the target sequences. SNP genotypes were assigned to target sites with at least 10X sequence coverage. Sites with <15% alternate allele frequency were considered homozygous genotypes.

Genetic analyses

Population structure was determined using Bayesian Markov chain Monte Carlo (MCMC) clustering model using the program STRUCTURE v2.3.4 (25, 26). The admixture model using sampling locations as prior was used with a burn-in of $2x10^5$ and $2x10^5$ iterations. There were 20 independent runs for each K value from 1 to 10. Arlequin was used to calculated population pairwise F_{ST} within each STRUCTURE hypothesized population and GenAlEx was used to generate a principle coordinates analysis of the clusters (28-30).

Phylogenetic analyses were conducted with MEGA6 using the Maximum Likelihood method based on the General Time Reversible model (31). The bootstrap consensus tree was inferred from 1000 replicates. Evolutionary rate differences among sites were modeled with a discrete Gamma distribution with the rate variation model allowing for some sites to be evolutionarily invariable (G + I). Initial trees for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach.

Results

Genotypes

Genomic DNA from 375 samples, including 33 samples from a historical panel of race types, was submitted for targeted sequencing. Of these samples, 167 field samples and 26 of the historical race types amplified with at least 20x coverage to be included in further analysis. The 167 field samples were 68 samples from Yuma, Arizona and 99 from Salinas Valley, California. The 26 race types included races 3, 4, and 10-14. Forty seven markers had a minimum of 10X sequence coverage for all 193 samples. One marker was fixed for homozygosity and removed and the remaining 46 markers were used for subsequent analysis.

One set of 96 samples was sequenced on both the MiSeq and the Ion Proton platforms. Results from each platform were identical. The remainder of the samples were sequenced on the Ion Proton.

In total, there were 82 unique multi-locus genotypes with 21 unique mutli-locus genotypes found more than once and assigned a genotype identifier of G1-G21 (Table 3). The alternate allele frequency in the collection of 82 unique genotypes ranged from 1 to 49% and averaged 27%, with the majority (74%) of the loci having alternate allele frequencies >20% (table 6, appendix). Genotype G1, recovered from 27 samples or 14% of the total analyzed samples, was found only in Yuma, Arizona on both organic and non-organic spinach cultivars Coati, Tasman, Carmel, Meerkat, Dramaderi, and Merlat. The second most common genotype, G2, was only recovered from Salinas Valley, California on the spinach cultivars Coati, Tasman, Carmel, Meetkat, Soloman, Platypus, Silverwhale, and PV1030. The third most frequent genotype, G3, was recovered from 17 samples from fields in Yuma and the Salinas Valley, specifically from cultivars Coati, Tasman, Silverwhale, Carmel, Banjo, Meerkat, and NH185. The G3 genotype also included a soon to be described new race (race 15) and is the only race-typed (or soon to be race-typed) sample that clustered with the majority of the field isolates. While most identical genotypes were recovered from field samples, there were two cases, G13 and G14, where identical genotypes were recovered for two race 10 (G13) and two race 13 (G14) samples (Table 3).

Population structure

Analysis of the program STRUCTURE indicated a K value of 6 based on the largest ln Pr(X|K) value. Figure 6 is a structure graph showing the genotypes divided by location and assigned a color based on cluster. The first cluster includes 11 genotypes from both Yuma and Salinas Valley. The second cluster includes 13 races from our panel including races 4, and 10-14. The third cluster has 10 unique genotypes and includes mostly samples from Yuma and one from Salinas Valley. In the fourth cluster, races 3, 10-12, and 14 are present. The fifth cluster is the largest cluster with 34 unique genotypes. These samples are field samples from both Yuma and Salinas Valley collected from ten different spinach cultivars. The sixth cluster is the smallest with only three unique genotypes from California, all recovered from the spinach cultivar Calisto.

The maximum likelihood tree complemented the six structure groups described above (Figure 7). Structure clusters 2, 5, and 6 formed monophyletic groups whereas clusters 1, 3, and 4 formed polyphyletic groups. A principle coordinates analysis showed separation between structure clusters (Figure 8).

Discussion

This is the first fine-scale assessment of population structure for *Peronospora farinosa* f. sp. *spinaciae* using SNP markers. A robust panel of SNP markers were developed from *Peronospora farinosa* f. sp. *spinaciae*. The inclusion of a historical panel of isolates that have been assigned race types provided useful insight as to the utility of the markers included in the final analyses and indicates there is a reasonably high level of overall polymorphism.

Genotypes from field samples collected in Yuma and Salinas in 2014 mostly differ from the historical panel of race types. Identical genotypes can be found in both Yuma, Arizona and Salinas Valley, California and very closely related genotypes from both locations grouped in both structure and phylogenetic analysis. Most of the samples from

Table 3. Identical Genotypes of Downy Mildew

Table of identical genotypes. The last column is either spinach varieties or race-typed individuals. Race-typed individuals are in bold.

Genotype Identifier	# of samples	Location/Race	Spinach Varieties and *race-typed individuals*
G1	27	Yuma	Coati, Tasman, Carmel, Meerkat, Dramaderi, Merlat
G2	21	Salinas	Soloman, PV1030, Tasman, Silverwhale, Carmel, Meerkat, Coati, Platypus
G3	17	Yuma/Salinas	***1014APLP***, NH185, Tasman, Silverwhale, Carmel, Banjo, Meerkat, Coati
G4	12	Salinas	PV1030, Tasman, Silverwhale, Carmel, Meerkat, Coati, Platypus
G5	8	Salinas	Soloman, Silverwhale, Meerkat, Coati, Platypus
G6	7	Yuma	Coati, Merlat
G7	6	Yuma/Salinas	NH185, Plover, Cello, Dramaderi
G8	4	Yuma	Cello, Dramaderi
G9	3	Yuma	Plover, Meerkat, Merlat
G10	3	Salinas	Soloman, Platypus, Silverwhale
G11	3	Yuma	Tasman
G12	3	Salinas	Silverwhale, Soloman, Coati
G13	2	10	***UA2509A, INT-1***
G14	2	13	***UA2411, UA1411A***
G15	2	Salinas	Calisto
G16	2	Yuma/Salinas	Gazelle, Dramaderi
G17	2	Yuma	Molakai
G18	2	Salinas	Meerkat, Silverwhale
G19	2	Salinas	Silverwhale, Meerkat
G20	2	Salinas	Meerkat, Coati
G21	2	Salinas	Banjo, Meerkat

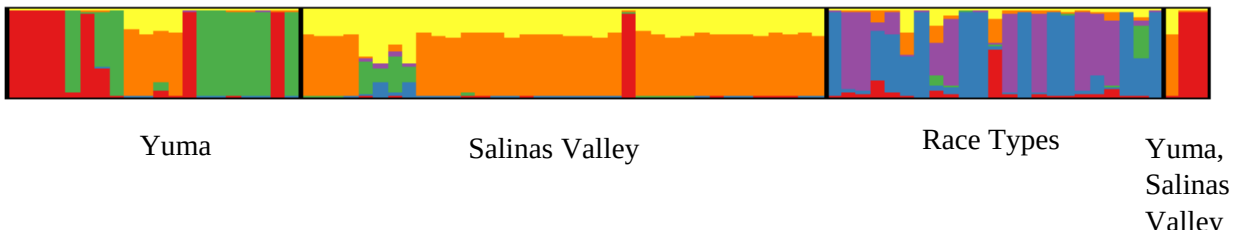
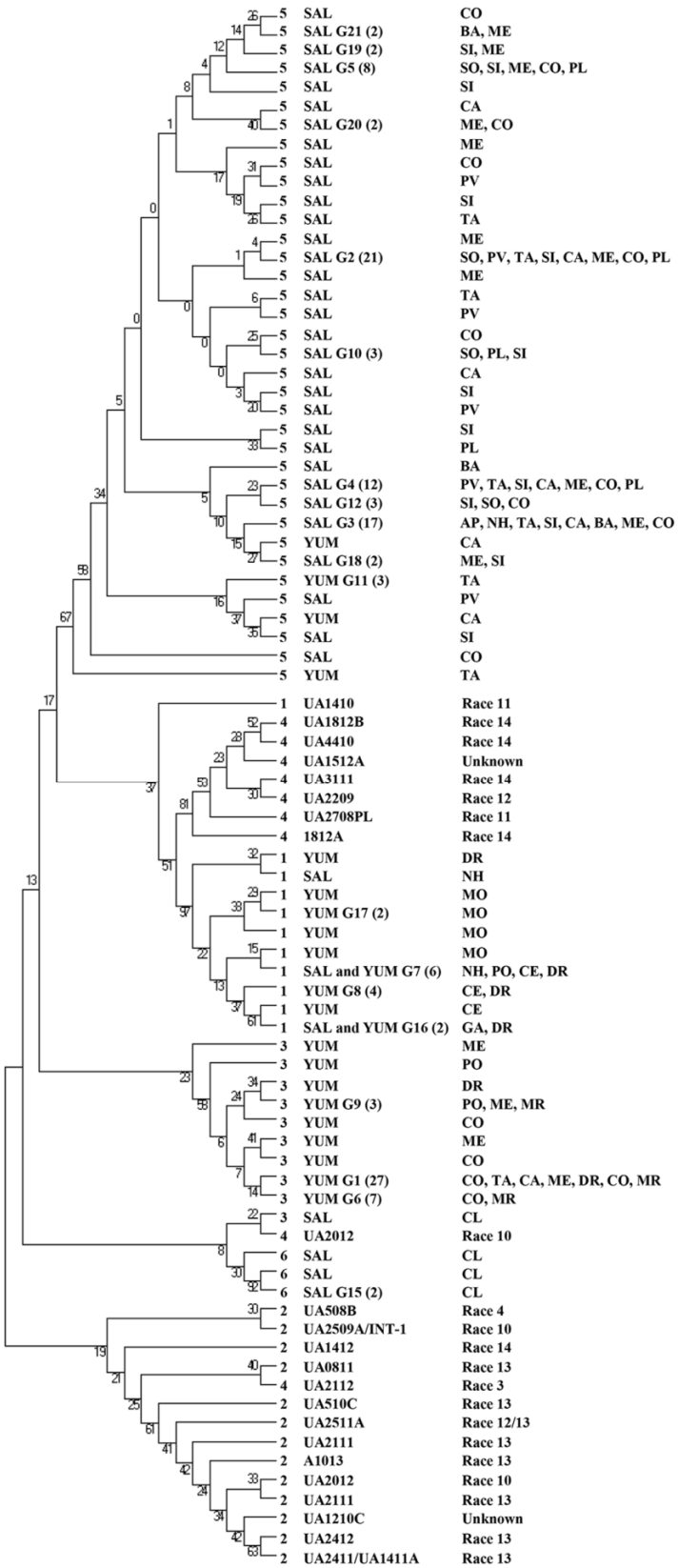


Figure 6. Structure Plot for Downy Mildew

STRUCTURE plot of assigned probabilities of 82 unique genotypes. Each genotype is represented by a bar indicating likelihood of membership in cluster 1 (red), cluster 2 (blue), cluster 3 (green), cluster 4 (purple), cluster 5 (orange), cluster 6 (yellow).

Figure 7. Maximum Likelihood Tree for Downy Mildew

Maximum likelihood tree for *Peronospora farinosa f. sp. spinaciae* genotypes. Three columns of data at each terminal branch include: STRUCTURE group 1-6, race ID or location (YUM = Yuma, AZ, SAL = Salinas Valley, CA) and genotype identifier with number of identical genotypes, and spinach varieties from which genotype was recovered (key on right of tree).



Spinach Cultivar
CO = Coati
SI = Silverwhale
CA = Carmel
ME = Meerkat
PV = PV1030
TA = Tasman
PL = Platypus
BA = Banjo
DR = Dramaderi
NH = NH185
MO = Molakai
CE = Cello
PO = Plover
CL = Calisto
SO = Soloman
AP = 1014APLP
GA = Gazelle
MR = Merlat

Figure 7. Continued

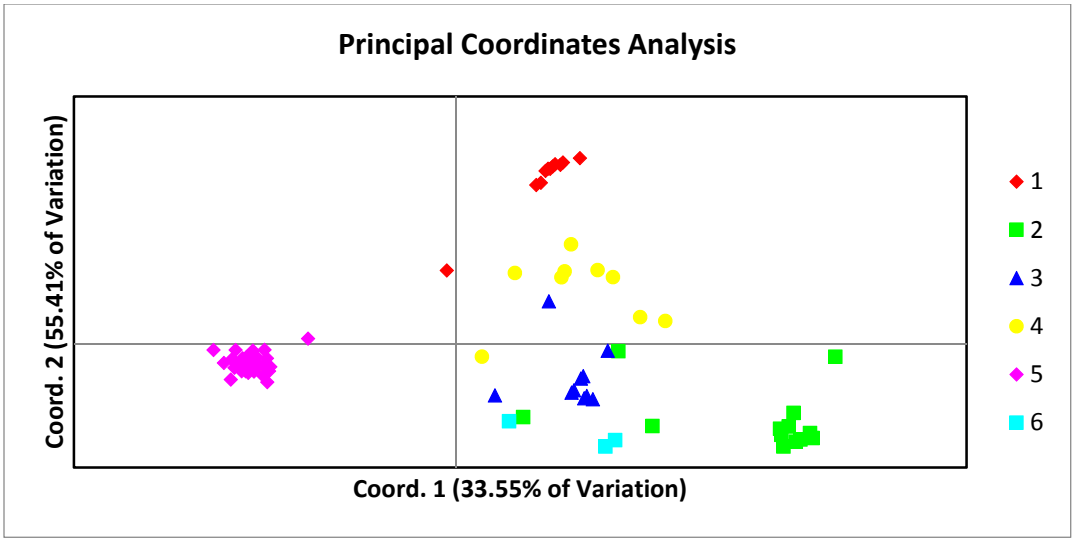


Figure 8. Principle Coordinates Analysis (PCoA) for Downy Mildew
Principle Coordinates Analysis (PCoA) showing genetic distances between individuals in six clusters. Each color represents a STRUCTURE cluster. Cluster 1 (red), cluster 2 (green), cluster 3 (blue), cluster 4 (yellow), cluster 5 (pink), and cluster 6 (teal).

these two production areas in 2014 had very similar genotypes. This suggests there is either movement between the two production regions during the concurrent growing seasons or there is a common source of downy mildew.

Isolate UA1014 APLP was the only isolate from the race panel that was found in group five, the largest group of field samples. This isolate is thought to be a new race and is currently in a ring test with labs around the world to be nominated as race 16. Closely related genotypes to this isolate dominated the field populations from 2014. The appearance of this new race and the many genotypes similar to it, is likely due to the selection pressure imposed by growers who have increasingly relied on spinach cultivars resistant to races 1-15 in Yuma and Salinas Valley.

Although some historical known races cluster together (e.g. Race 14); there is not a strict correlation between genotype and race type. Therefore, it is not possible with this panel of neutral SNP markers to strictly assign a race type to an unknown isolate based solely on the multi-locus SNP genotype.

It is becoming increasingly clear that oomycete plant pathogens have the potential to produce significant amounts of genotypic variation during asexual population expansion (51). The majority of the genotypes recovered during the 2014 season are very similar and due to the number of heterozygous sites in common, it is likely the differences observed are due to asexual phenomenon, such as loss of heterozygosity. Current studies that are in progress, based on whole genome re-sequencing, indicate loss of heterozygosity can play an important role in generating asexual variation for spinach downy mildew. Although the 2014 epidemics appear to be dominated by a few clonal lineages with sub-clonal variation, there appears to be sufficient diversity in the form of unique multi-locus genotypes in the historical race panel to suggest sexual as well as asexual mechanisms contribute to the overall diversity over time. Further studies over time to track genotypic diversity in parallel with the deployment of novel spinach cultivars will be helpful to better understand how this important pathogen is able to rapidly overcome novel resistant plant varieties and may be useful to determine the importance, overall, in sexual vs. asexual mechanisms for driving variation in epidemic populations.

Conclusions

Advances in genetic technology are giving an unprecedented view of plant pathogen populations and their evolution in field scenarios. Oomycetes such as those studied here are notoriously genetically plastic and can quickly lose “abilities” (e.g. spore production, virulence, etc.) when manipulated in a laboratory or greenhouse setting. In both of these studies, unlike the majority of previous work on oomycetes in the field, genotypic variation within a panel of pathogen genes was measured directly from infected plant tissue. Similar to most oomycete plant pathogens, P. infestans and P. farinosa produce millions of asexually-derived spores on the surface of infected tissues. The lesions are typically exposed to intense sunlight and the fidelity with which the genomes of these spores are mitotically replicated is unknown. These studies suggest that within the context of an epidemic, the fidelity is not 100% and that genotypic diversity is being produced asexually. Much more work is needed to understand the dynamics of how this variation is produced, the overall frequency, and most importantly, the evolutionary impact.

References

1. Lebeda A. The Downy Mildews - Genetics, Molecular Biology and Control. Cooke BM, Spencer-Phillips PTN, SpringerLink, editors. Dordrecht: Dordrecht : Springer Netherlands; 2008.
2. Fry WE, Grünwald NJ. Introduction to Oomycetes. The Plant Health Instructor. 2010.
3. Lamour K. Phytophthora : a global perspective. Cambridge, MA: Cambridge, MA : CAB International; 2013.
4. Cárdenas Toquica M, Restrepo Restrepo S. Phytophthora infestans population structure: a worldwide scale. Acta Biológica Colombiana. 2012;17(2):227-40.
5. Adler NE, Erselius LJ, Chacon MG, Flier WG, Ordonez ME, Kroon L, et al. Genetic diversity of *Phytophthora infestans* sensu lato in Ecuador provides new insight into the origin of this important plant pathogen. Phytopathology. 2004;94(2):154-62.
6. Abad ZG, Abad JA. Another look at the origin of late blight of potatoes, tomatoes, and pear melon in the Andes of South America. PLANT DISEASE. 1997;81(6):682-8.
7. Correll JC, Bluhm BH, Feng C, Lamour K, du Toit LJ, Koike ST. Spinach: better management of downy mildew and white rust through genomics. Eur J Plant Pathol. 2011;129(2):193-205.
8. Irish BM, Correll JC, Koike ST, Morelock TE. Three new races of the spinach downy mildew pathogen identified by a modified set of spinach differentials. PLANT DIS. 2007;91(11):1392-6.
9. Choi YJ, Thines M, Han JG, Shin HD. Mitochondrial phylogeny reveals intraspecific variation in Peronospora effusa, the spinach downy mildew pathogen. Journal of microbiology (Seoul, Korea). 2011;49(6):1039-43.
10. Fry W. *Phytophthora infestans*: The plant (and R gene) destroyer. 2008. p. 385-402.
11. Cardenas M, Grajales A, Sierra R, Rojas A, Gonzalez-Almario A, Vargas A, et al. Genetic diversity of *Phytophthora infestans* in the Northern Andean region. BMC GENETICS. 2011;12.
12. Adler NE, Chacón G, Flier WG, Forbes GA. The Andean fruit crop, pear melon (*Solanum muricatum*) is a common host for A1 and A2 strains of *Phytophthora infestans* in Ecuador. Plant Pathology. 2002;51(6):802.
13. Chacón MG, Adler NE, Jarrin F, Flier WG, Gessler C, Forbes GA. Genetic structure of the population of *Phytophthora infestans* attacking *Solanum ochranthum* in the highlands of Ecuador. European Journal of Plant Pathology. 2006;115(2):235-45.
14. Vargas AM, Quesada Ocampo LM, Céspedes MC, Carreño N, González A, Rojas A, et al. Characterization of *Phytophthora infestans* populations in Colombia: First report of the A2 mating type. Phytopathology. 2009;99(1):82-8.
15. Forbes GA, Escobar XC, Ayala CC, Revelo J, Ordoñez ME, Fry BA, et al. Population genetic structure of *Phytophthora infestans* in Ecuador. Phytopathology. 1997;87(4):375-80.
16. Delgado RA, Monteros-Altamirano AR, Li Y, Visser RGF, van der Lee TAJ, Vosman B. Large subclonal variation in *Phytophthora infestans* populations associated with Ecuadorian potato landraces. Plant Pathol. 2013;62(5):1081-8.
17. Erselius LJ HH, Ordoñez ME, Oyarzun PJ, Jarrin F, Velasco A, Ramon MP, Forbes GA. Genetic diversity among isolates of *Phytophthora infestans* from various hosts in Ecuador Lima, Peru: International Potato Center1999.
18. Oyarzun PJ, Pozo A, Ordonez ME, Doucett K, Forbes GA. Host specificity of *Phytophthora infestans* on tomato and potato in Ecuador. Phytopathology. 1998;88(3):265-71.
19. Oliva RF, Erselius LJ, Adler NE, Forbes GA. Potential of sexual reproduction among host- adapted populations of *Phytophthora infestans* sensu lato in Ecuador. Plant Pathol. 2002;51(6):710-9.
20. Abbott CL, Gilmore SR, Lewis CT, Chapados JT, Peters RD, Platt HW, et al. Development of a SNP genetic marker system based on variation in microsatellite flanking regions of *Phytophthora infestans*. Canadian Journal of Plant Pathology. 2010;32(4):440-57.

21. Lamour K, Hu J. Diversity and Phytophthora: a threat to forests, crops and traditional laboratory research - Mini Review. CAB Reviews. 2013;8:doi: 10.1079/PAVSNNR20138038.
22. Haas B, Kamoun S, Zody MC, Jiang R, Handsaker R, Cano L, et al. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature. 2009;461(7262):393-8.
23. Yoshida K, Schuenemann VJ, Cano LM, Pais M, Mishra B, Sharma R, et al. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. eLife. 2013;2.
24. Nguyen-Dumont T, Pope BJ, Hammet F, Southey MC, Park DJ. A high-plex PCR approach for massively parallel sequencing. BioTechniques. 2013;55(2):69-74.
25. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945-59.
26. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes. 2004;4(1):137-8.
27. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 2012;4(2):359-61.
28. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10(3):564-7.
29. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 2012;28(19):2537-9.
30. Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes. 2006;6(1):288-95.
31. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725-9.
32. Catal M, King L, Tumbalam P, Wiriyajitsomboon P, Kirk WW, Adams GC. Heterokaryotic nuclear conditions and a heterogeneous nuclear population are observed by flow cytometry in *Phytophthora infestans*. Cytometry Part A. 2010;77(8):769-75.
33. Van Der Lee T, Testa A, Robold A, Van, t Klooster J, Govers F. High- density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics. 2004;167(4):1643-61.
34. Morgan WK, Sophien. RXLR effectors of plant pathogenic oomycetes. Current Opinion in Microbiology. 2007;10(4):332-8.
35. Whittaker SL, Shattock RC, Shaw DS. Variation in DNA content of nuclei of *Phytophthora infestans* as measured by a microfluorimetric method using the fluorochrome DAPI. Mycological Research. 1991;95(5):602-10.
36. Brandenberger LP, Morelock TE, Correll JC. Evaluation of Spinach Germplasm for Resistance to a New Race (Race 4) of Peronospora farinosa f. sp. spinaciae. Hortscience. 1992;27(10):1118-9.
37. Correll JC, Morelock TE, Black MC, Koike ST, Brandenberger LP, Dainello FJ. Economically Important Diseases of Spinach. PLANT DISEASE. 1994;78(7):653-60.
38. Choi YJ, Hong SB, Shin HD. Re- consideration of Peronospora farinosa infecting Spinacia oleracea as distinct species, Peronospora effusa. Mycological Research. 2007;111(4):381-91.
39. Food and Agriculture Organization of the United Nations: FAOSTAT 2014.
40. Feng C, Mansouri S, Bluhm BH, du Toit LJ, Correll JC. Multiplex real- time PCR assays for detection of four seedborne spinach pathogens. Journal of Applied Microbiology. 2014.
41. Lucier G. Factors affecting spinach consumption in the United States. In: United States. Dept. of Agriculture. Economic Research S, United States. Dept. of Agriculture. Office of Economics. Economic Research S, E.R.S, editors. Electronic outlook report from the Economic Research Service. Washington, D.C.]: Washington, D.C. : U.S. Dept. of Agriculture; 2004.

42. Foss CRJ, L. J. . Crop Profile for Spinach Seed in Washington. 2005.
43. Inaba T. Seed Transmission of Spinach Downy Mildew. Plant Disease. 1983;67(10):1139.
44. Correll JC, Koike ST, Brandenberger LP, Black MC, Morelock TE. A New Race of Downy Mildew Threatens Spinach. California Agriculture. 1990;44(6):14-5.
45. Irish BM, Correll JC, Koike ST, Schafer J, Morelock TE. Identification and cultivar reaction to three new races of the spinach downy mildew pathogen from the United States and Europe. PLANT DIS. 2003;87(5):567-72.
46. Feng C, Correll JC, Kammeijer KE, Koike ST. Identification of new races and deviating strains of the spinach downy mildew pathogen Peronospora farinosa f. sp. spinaciae. Plant Disease. 2014;98(1):145-52.
47. Eenink A. Linkage in Spinacia oleracea L. of two race- specific genes for resistance to downy mildew Peronospora farinosa f. sp. spinaciae B yford. International Journal of Plant Breeding. 1976;25(1):713-5.
48. Brandenberger LP, Correll JC, Morelock TE. Identification of and Cultivar Reactions to a New Race (Race- 4) of Peronospora farinosa f. sp. Spinaciae on Spinach in the United States. PLANT DIS. 1991;75(6):630-4.
49. Satou M, Nishi K, Kubota M, Fukami M, Tsuji H, Ettekoven KV. Appearance of race Pfs: 5 of spinach downy mildew fungus, Peronospora farinosa f. sp. spinaciae, in Japan. Journal of General Plant Pathology. 2006;72(3):193-4.
50. Lamour K, Finley L. A strategy for recovering high quality genomic DNA from a large number of *Phytophthora* isolates. Mycologia. 2006;98(3):514-7.
51. Hu J, Diao Y, Zhou Y, Lin D, Bi Y, Pang Z, et al. Loss of Heterozygosity Drives Clonal Diversity of Phytophthora capsici in China. Plos One. 2013;8(12).

Appendix

Table 4. *P. infestans* primers

Name	Primer
1.1_2422620_F1	CCACCCGCATGGCGTCGTT
1.1_2422620_R1	CGCTAATGTCTTTGTTGTCCTC
1.1_3023095_F1	ACACCGTTCTGGTTGTTGTCG
1.1_3023095_R1	GTAGGTCCTCATTTTCCCTTTTT
1.1_3547675_F1	CTGTGGGCTCACAGTTTAAACA
1.1_3547675_R1	GTCGATCAAAGTGAAGGCAGC
1.1_5516055_F1	CGAGCCCAGCGACACGTAC
1.1_5516055_R1	GTCATCGTGCAGCGCACCG
1.1_931590_F1	AGATCATTCGAAGTCAGAGTGG
1.1_931590_R1	CGGCGACAAGTCTTCTTCATC
1.1020399_F1	GGCAATGAGTGTGAGACCGTT
1.1020399_R1	ACGCTCGACTACATTCGCGG
1.10_651437_F1	GAATCCGCGCGCTGTAACCA
1.10_651437_R1	CAATCCGGACAAACGTCGTGT
1.101_224249_F1	TAGGTTTGCTGAGTGACGAATC
1.101_224249_R1	AGTCCACGATCATGTGGTAATC
1.11535743_F1	CGGAGAACTAGGGACTCAGAA
1.11535743_R1	CAGATATGCCGTTCTGTAGCC
1.11730516_F1	CGCAAACGATGATGCAATGGC
1.11730516_R1	ATGCTTTTTCATATACTGGACAAG
1.11786659_F1	TTCGAGCAAGTAAACTCAATGTTT
1.11786659_R1	TAGCCGACATCGTCGCTTCTT
1.11_2708319_F1	AAGCTGTCGACTATGTGCACTA
1.11_2708319_R1	GAGACCAGACCGTAGTAGAAG
1.11_3350688_F1	CTTCTCCGTCTCTAATTCCGC
1.11_3350688_R1	AAGCGGAATTCTATAGTGAGAGA
1.11_360197_F1	GAGTTGAGCAAACACCCGAAC
1.11_360197_R1	TAGGTCATCATATACACTCTCCA
1.11028637_F1	CATAATTGGCGCCTGTGGCG
1.11028637_R1	CGAATCTAATGCCGAATAGTAAG
1.12773750_F1	GGATGACGTGTACAAAGCTAAC
1.12773750_R1	TCGTACAAACAGAACTTGAGGTT
1.13_2911071_F1	ATAACTTTTGTGCTCTTTTTTGCTT
1.13_2911071_R1	CTCCGAGAGATTCTGCTGGAA
1.13108866_F1	CTTGCGAGTGGGTACGTGAG
1.13108866_R1	GAACGATGTGCAGGACGCCA
1.13228908_F1	GACTACTGGGACAGGATTCTAT
1.13228908_R1	GCTGGTAGCACCCAGAGCC

Table 4. Continued

Name	Primer
1.14462417_F1	AAACTGTGAGATGCGATGCGC
1.14462417_R1	CGACGTCCTTCAGCGACGC
1.14932533_F1	AAGTCGGCACAAGTGAAGACG
1.14932533_R1	GGTTTCGAACGATCTCAGCTG
1.14_2071843_F1	TTTTATGCATCAGACCGTCGAC
1.14_2071843_R1	GCTTACAGGCCGACGATGAC
1.15_2282336_F1	CACGTGGTGCGGTCCGTGT
1.15_2282336_R1	TCCACGCATCTCGTCGCACT
1.15_2332768_F1	AGCAGTTTCTTCGCGGCCAC
1.15_2332768_R1	ATGACGCGTTTACTGCAAGCAA
1.16_588934_F1	GCACACTCAAAGGAAAGCACTT
1.16_588934_R1	AGGGGTTGTGGTCTTATGGCA
1.17706137_F1	GTACAGCACCATAAAGACACTC
1.17706137_R1	GTCGTGGATTGGATTAAAGCGT
1.17717761_F1	CCCAGATTGTCCAGTCCCAC
1.17717761_R1	ACGTCGGAAATATGCTTACTGTA
1.17_887082_F1	GGGTAGTACTGAGGCGAGTG
1.17_887082_R1	TCACTCTGCTGAATGTGTTCCA
1.18156420_F1	GGCACCAAACGAGAAACCAGT
1.18156420_R1	CAGCTTCGGATCAACTACGTC
1.18_835118_F1	TTCGCCTTGCCATTGTTTAAGC
1.18_835118_R1	GTTGGAAGTACTGGATGATTACA
1.19085064_F1	GGGTCGCTCAGCTGCGAAAA
1.19085064_R1	CGTGGGGGACATTATCAAGGT
1.19425123_F1	ATGTAGCTACCATGCATCACAC
1.19425123_R1	TCGCAGCGAATGCCACCTGT
1.2316873_F1	ACGCGAAGCAGCCTCCACG
1.2316873_R1	ACCGACTCAGGTGTACGCTG
1.2_3934005_F1	ATCAATTCAAGAGAATAAAAGACGA
1.2_3934005_R1	CGTTACAGGGATCTTTATAAACG
1.2_4048575_F1	TGCTGTACGGATTGTTCTAAGAT
1.2_4048575_R1	TCCGCGTGTCGATTTGTCTGA
1.20_436558_F1	AGTCCGTCATTGCCAAGGTTG
1.20_436558_R1	TCTTGCACGTTGTCACTCAGC
1.21258416_F1	ATATCGTTAACTGCGTCATCAAC
1.21258416_R1	CTGTGGCTGGCCATGCCAC
1.21_544320_F1	ACAGCCACCAGTGTCGCGAT
1.21_544320_R1	AAATGATCCGAACCAGATTTAGG
1.25192926_F1	TCAGCTACTGTACGTCCATCC

Table 4. Continued

Name	Primer
1.25192926_R1	CCAGCAGCAGTGTCTGTTGC
1.25633422_F1	TGGGTCCGCAATGTGTTATGG
1.25633422_R1	CGCCAGTGCTCGCTTTACTC
1.25_876700_F1	TGGTGCCATCTCGGAAGGGT
1.25_876700_R1	AAAGGCAGCGGGTAGTGCAC
1.26_210006_F1	CGCGGGCAAAGATACCGCTT
1.26_210006_R1	GCAACATGGTCGTCGTGGAG
1.27_816888_F1	ACGACTCTCGCGTGCGGGT
1.27_816888_R1	AATATTCGATCGATTTCAGATGAG
1.27_852653_F1	GTCGGGGCCGCATAAAATCC
1.27_852653_R1	ACCTTAATCGCAGCCGTCACA
1.28695210_F1	TCTGGATGAACGGCCTCGAG
1.28695210_R1	TTCGTCATACATATCCAAATCCC
1.29_581789_F1	ACGTCGGCTGCTGTCGAGG
1.29_581789_R1	CCTAACGAAACAAGAGCCTCG
1.29_852731_F1	TACGATCTCAAACACTGAAAACAT
1.29_852731_R1	TGAGACAGTGAGTAGCTTGATC
1.3_2243661_F1	GTGGGAACTGCGTCGTCTGT
1.3_2243661_R1	GCTGGCACGCAGCACGATC
1.3_3220722_F1	AGCGACTGCAGTTCCTTGAATA
1.3_3220722_R1	AACAAGCTGCGTGCTGCAAGT
1.3_4145204_F1	TCTTCCCATTCAAGCATGGGTT
1.3_4145204_R1	CGGGATACCGTGGCGTCGA
1.30_440315_F1	ACTCTGCTACTGTTTCTCAGGA
1.30_440315_R1	AGAAGCACATGCCGTCCAGG
1.302_59924_F1	CAGTCCAAGACGGTGTTGCG
1.302_59924_R1	CATGACACTCTTAACCGGCAC
1.315_63639_F1	CCTCAGGATACCTGGCCGAT
1.315_63639_R1	GTGCCCTCGATCAGCAAATAG
1.32_896666_F1	GTACTTTTTGCCTTCGATGATCT
1.32_896666_R1	ACCTGCACTGCTTGCGTCCT
1.33_979119_F1	TCATTTCTGATTCCCGTTCTATG
1.33_979119_R1	ACGATGAAGACGACGATGAAGA
1.34_566592_F1	AGACTTCTTCGCAAGAACGAAG
1.34_566592_R1	TGCTGCCACCACATGTATGCA
1.34_566663_F1	CGAATCTAAATGAGGAGATGTTTA
1.34_566663_R1	TGCGCCTTGCGTCTTGCGG
1.34_578357_F1	CAGCTCACTACATACACGATTC
1.34_578357_R1	ATCTTATCATTACCCATGAGCTG

Table 4. Continued

Name	Primer
1.34_951463_F1	GCAGTACTGCCGAAACTCTCA
1.34_951463_R1	GTACGCAAGTACCGCGGCAA
1.36_905326_F1	ATCGCTGTTTCGAGGGCTGAA
1.36_905326_R1	ACAGCGATTTTGTTGCGTCAAC
1.4_2458623_F1	TTGAAAAGCGACTTGTGAAACAAA
1.4_2458623_R1	CGGCCGGTGGTCTTGTTGC
1.4_3096042_F1	TCGACTCGCCAATCTGTACAC
1.4_3096042_R1	CACAATATGAGCCGGTCTGGA
1.40_751211_F1	TCCGTCTGCATCGAGCTTACA
1.40_751211_R1	GCAACGACCACCGAAAATGTG
1.40_977746_F1	ATTTGATCTTCGACTTTTTGTAACT
1.40_977746_R1	GTGGGGAAGACGAACAAGATG
1.43_780367_F1	CGTTCTTCATCACCTTCTCGG
1.43_780367_R1	GCTAGCAGCGGCCATTGTGT
1.44_558022_F1	AGGTCAAAGAGAAAGGGAAACC
1.44_558022_R1	CTCCCTTGGACCTGGGAGG
1.5367534_F1	AGCAGGGTGTTGCTTATGTCAA
1.5367534_R1	GCAAGATCGTCGAAAATTTTCTC
1.5450337_F1	GAGTCGCAGGCCGGGAAGA
1.5450337_R1	AGGCCTCTTTCTCTTTTGCAGT
1.5_3518729_F1	TATATTCTTCTGCATTCTTGAGGA
1.5_3518729_R1	CTGAACACCTCGTTCGCACG
1.50_304997_F1	ACTCCATCACCACAGCGTGC
1.50_304997_R1	ACGAGCGTCAAGAACTCCCC
1.50_432212_F1	ATGACGGCGTCCAGCTTCTG
1.50_432212_R1	GCGTGGCCATTGTGCACGC
1.51_910750_F1	CTCCGTAGTTTTGGATGACGTA
1.51_910750_R1	GAATTGAATCTGAGTGCCCACA
1.53_801138_F1	AGCGGTGAACCCCTTCCAGA
1.53_801138_R1	CTCATGACGCCGTTTAGTCCT
1.6_2277226_F1	GCAAGCGTGGCCTTGCTTCA
1.6_2277226_R1	GCTCAGTGGACGTCGACATC
1.60_313677_F1	CTGTAGCAGTCGCTTGGCTG
1.60_313677_R1	GCGCAGAATTCGCGTCTGGA
1.63_646558_F1	TTCTAATTGTCTATTGGAAGGGTT
1.63_646558_R1	ACGAGGAGTGAAGCGCAGTG
1.64_470447_F1	GAAGAACCGAGAGGAACTGCA
1.64_470447_R1	TCTGCGCCGACTGCAGCTC
1.66_466001_F1	TTTGTCTCGTCGAGTTTTGCAG

Table 4. Continued

Name	Primer
1.66_466001_R1	TGGTCCAGACAATAGAAAAACTG
1.67_337817_F1	CTTGCGCAGATTCCACGCTG
1.67_337817_R1	TTGACGGCTCCGCTCACAGA
1.67_362687_F1	TTCAACGACGCAAATTTCGGAC
1.67_362687_R1	GATGTTGTCCAGGTCCGTGG
1.69_251262_F1	TAAGCCACAGAAGCGCACACT
1.69_251262_R1	TCTCAAGGGTTTTCTTGCCGAT
1.7_3067555_F1	CGAGACACAGCACGAGTGTC
1.7_3067555_R1	TGTGGCTCACGTCCGTGGC
1.7_3487748_F1	ATCACATTGCAGCTGTGACAGT
1.7_3487748_R1	ACCCGAATGTGCTGCGCATG
1.7_4041958_F1	TCGACATCAAGGGATCCTGGA
1.7_4041958_R1	ACGGATGCTGTACTCACAGCT
1.72_552467_F1	CCTCCTTGACATCCGAATCGT
1.72_552467_R1	TCCGGTACCCATGCTATCTCT
1.8767503_F1	GTTCTCAGGCCACTCGAGAG
1.8767503_R1	GTTCTCGAAAGTGTATCCAGCA
1.8_2653523_F1	ATCTCTTCCGGTGTTTTGCCTT
1.8_2653523_R1	ATGACGCAGTTGATGTGGACG
1.8_2944433_F1	CCCGAGGCCTCTGTGATGC
1.8_2944433_R1	AGCGTTTAAACATGCTGAAACAC
1.83_45429_F1	ACGACGACGAGAGGCCACAT
1.83_45429_R1	TCTTCCTCATCGTCTCCTACG
1.89_310701_F1	GGTCTCAAGCACGAAGAAGCT
1.89_310701_R1	TGTTCAATTCATCGCTTTCGTTAT
1.9295185_F1	TACGCTTTTACACTTGCTGCTAT
1.9295185_R1	ACACACGCGTGCTTGAATGAAT
1.9_2581958_F1	GGTGACCACTGTCGAAAACGT
1.9_2581958_R1	ACAGTCAGCGATGTCAATGGC
1.9_3123848_F1	GCCCAAGCCAATTACCGAGAA
1.9_3123848_R1	CTTCTTTTGCCTTTTTCACGGC
1.9_3721006_F1	AGTGCCTCAAACAATGAGTTTCT
1.9_3721006_R1	AGGAAGGCTGGCAGAGTCATT
1.9_3795561_F1	TTTGCAGAAGCTTTGAGCAACG
1.9_3795561_R1	CAATGACGATAGAGACAGAGCA
1.9537980_F1	TTGCGTCGCTCCTACGAAGAA
1.9537980_R1	AACGGCTGACTCGCCTTATAC

Table 5. *P. infestans* information on SNP locations

Supertercontig_position	Allele	Gene ID	Annotation	Effect	AA change
1.1_3023095	A/C	PITG_00582	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	L/R
1.1_3547675	A/C	PITG_00676	transcription factor	SYNONYMOUS_CODING	P/P
1.1_5516055	A/T	PITG_00921	phospholipase?D,Pi-PLD-like-1	NON_SYNONYMOUS_CODING	Y/N
1.1_931590	A/T	PITG_00140	chromodomain protein, putative	NON_SYNONYMOUS_CODING	S/T
1.10_120399	A/G	PITG_06715	serine/threonine-protein phosphatase PP2A catalytic subunit	SYNONYMOUS_CODING	D/D
1.101_224249	T/C	PITG_19671	N-acetyltransferase 10	NON_SYNONYMOUS_CODING	L/P
1.11_1535743	A/T	PITG_22824	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	N/I
1.11_2708319	T/C	PITG_07661	Major Facilitator Superfamily (MFS)	SYNONYMOUS_CODING	L/L
1.11_3350688	A/G	PITG_07744	DEAD/DEAH box RNA helicase, putative	NON_SYNONYMOUS_CODING	A/V
1.11_360197	T/C	PITG_07387	Avr4 secreted RxLR effector peptide	NON_SYNONYMOUS_CODING	L/S
1.110_128637	T/C	PITG_19846	hypothetical protein	STOP_GAINED	W/*
1.12_1773750	C/G	PITG_08032	P-type ATPase (P-ATPase) Superfamily	NON_SYNONYMOUS_CODING	Q/H
1.13_2911071	A/G	PITG_08599	thimet oligopeptidase	SYNONYMOUS_CODING	P/P
1.132_128908	A/G	PITG_20051	predicted protein	STOP_GAINED	W/*
1.14_1462417	A/C	PITG_09218	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	A/A
1.14_2071843	A/G	PITG_09316	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	S/F
1.17_1706137	T/C	PITG_10116	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	E/E
1.17_1717761	T/C	PITG_10119	ubiquitin-specific protease, putative	NON_SYNONYMOUS_CODING	G/E
1.18_1156420	A/G	PITG_10416	carbohydrate-binding protein, putative	SYNONYMOUS_CODING	A/A
1.19_1425123	T/C	PITG_11155	conserved hypothetical protein	SYNONYMOUS_CODING	Y/Y
1.2_1316873	T/C	PITG_01398	glycoside hydrolase, putative	NON_SYNONYMOUS_CODING	K/R
1.2_3934005	A/G	PITG_01724	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	G/S
1.2_4048575	A/G	PITG_01745	NEDD4-like ubiquitin ligase, putative	SYNONYMOUS_CODING	Y/Y
1.21_1258416	T/C	PITG_11953	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	D/G
1.25_1192926	T/C	PITG_13069	conserved hypothetical protein	STOP_GAINED	Q/*
1.25_1633422	A/G	PITG_13125	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	T/A

Table 5. Continued

Supertercontig_position	Allele	Gene ID	Annotation	Effect	AA change
1.25_876700	A/T	PITG_13031	vaiable flagellar number 3	NON_SYNONYMOUS_CODING	T/S
1.29_852731	T/C	PITG_13543	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	G/G
1.3_2243661	A/G	PITG_02479	T-complex protein 1 subunit theta	SYNONYMOUS_CODING	E/E
1.3_4145204	T/C	PITG_02821	cysteine protease family C13, putative	SYNONYMOUS_CODING	S/S
1.315_63639	T/C	PITG_21401	RNA methyltransferase	SYNONYMOUS_CODING	C/C
1.32_896666	A/G	PITG_14156	protein kinase	SYNONYMOUS_CODING	A/A
1.4_2458623	A/G	PITG_03627	cleavage induced conserved hypothetical protein (PITT_03627)	NON_SYNONYMOUS_CODING	L/S
1.4_3096042	T/G	PITG_03740	P-type ATPase (P-ATPase) Superfamily	SYNONYMOUS_CODING	T/T
1.40_977746	T/C	PITG_15728	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	E/K
1.43_780367	A/T	PITG_23069	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	I/F
1.44_558022	A/G	PITG_16464	Sulfate Permease (SulP) Family	NON_SYNONYMOUS_CODING	T/A
1.5_1367534	C/G	PITG_22729	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	Q/E
1.5_1450337	A/C	PITG_04145	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	H/Q
1.5_3518729	T/C	PITG_04367	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	G/G
1.51_910750	A/C	PITG_16738	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	I/S
1.53_801138	T/C	PITG_17348	Myosin protein	SYNONYMOUS_CODING	S/S
1.60_313677	T/G	PITG_17583	protein kinase, putative	SYNONYMOUS_CODING	I/I
1.64_470447	T/C	PITG_17936	conserved hypothetical protein	NON_SYNONYMOUS_CODING	F/L
1.69_251262	A/T	PITG_18510	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	E/D
1.72_552467	T/C	PITG_18405	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	T/T
1.8_2653523	T/C	PITG_22802	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	E/E
1.8_2944433	T/C	PITG_06570	conserved hypothetical protein	SYNONYMOUS_CODING	R/R
1.83_45429	T/C	PITG_19175	hypothetical protein	SYNONYMOUS_CODING	E/E
1.89_310701	A/C	PITG_19307	secreted RxLR effector peptide, putative	SYNONYMOUS_CODING	T/T
1.9_2581958	A/G	PITG_06004	ATP-binding Cassette (ABC) Superfamily	NON_SYNONYMOUS_CODING	T/A

Table 5. Continued

Supertercontig_position	Allele	Gene ID	Annotation	Effect	AA change
1.9_3123848	A/G	PITG_06077	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	E/K
1.9_3721006	T/C	PITG_06114	hypothetical protein	SYNONYMOUS_CODING	R/R
1.95_137980	A/C	PITG_19529	secreted RxLR effector peptide, putative	NON_SYNONYMOUS_CODING	M/L

Table 6. Downy Mildew SNP information

Contig_Position	Primers		Allele Counts for 82 unique genotypes				
SNP IDs	Forward	Reverse	A	C	G	T	% Alternative Allele
119_2816	GGCCTCCCTGGTAAATTGAT	AGTTGCAGCAAACGACACAG		163		1	**1**
934_15343	AGACTCCGTCCGAGTTTCAG	ATCCACAGGCGGAAATACAC	4		160		**2**
310_3511	GAGCGATTCCCTGTTGACTC	TGTTGCTCCTCATCCTGTGA	7		157		**4**
635_20522	CGAAGTGATCACCGTAACGA	GGAGTATCGGGGTTGGTACA		152		12	**7**
703_46170	CGCGTCTTGATATTGCTGAA	CAAGAAGGCACCATGTCAAA	15		149		**9**
288_15713	CAGTTGGTGGGAGGAAAATG	TCCTGAGGCCAAATCTCATC	18		146		**11**
87_6382	TGACGACTCCATAACGTCCA	ACCCTGCATCACAAGGAGAC			144	20	**12**
793_2335	GGCTCGCTCACACAACTTTT	CCAAAGTCAGCTTCGACCTC	139		25		**15**
1186_9551	ATGCTTTGAACTGCCGAGAT	TTCAAATTGACGGCGATAAA	138		26		**16**
1027_12305	CGGCTTCATCCTCCATGATA	CAGTGGAGGAGGACCTTTTG	27		137		**16**
148_17882	TGGCAACTTGGTTGAGGAAT	ACTGATTGCCCGCTTGAC		135		29	**18**
795_9577	CTTTAGCTGGGAGGTGTTGG	AGCACAGACGGGTTCGTAGT	134	30			**18**
1224_36034	ACAATAGCGGGTGCATTCAG	CTACTGAAACGCGGACCTGT	132		32		**20**
692_12307	TGATTGTTTTCGTCTCCTTCG	CGATGGACAGTCGCTGTAAC		132		32	**20**
76_6417	ACGCAGAGGCCTAAACAAAA	GTCACGATGGACCAGTTCAC		32		132	**20**
809_3531	CTGTTCTTCTTGTGCCAACG	GATCCGCAAGATGAAACACA	32		132		**20**
84_7881	GCGACGTGACAATCGTCTT	TGATTGTTCATGCGAGTGGT		131		33	**20**
234_4106	ATGTGGACTCGGCATCTCTT	GCATTGGATCGTTATGAAGC		128	36		**22**
689_10925	GTGGTGCTGTCTTTTGCTGA	GTTGCTGGTTGGTCTCCTTC	128		36		**22**
445_18260	CGCTATTGTTCGCGATGAT	CCAAAAAGAGCGGGTTTCTA	127		37		**23**
606_18824	AAACAATGGGTGTGGACCAT	GCTTTGATCGCACGTCATAC		39		125	**24**
536_39392	AGCACCATCGTCGTGCTAAT	ACACCCATGACCTCCAACAT		40		124	**24**
230_11674	GCGAAAAACTAGAGCCTGGA	ATACTGCTCGACCACGTCAA	123		41		**25**
597_1412	AGCCGAAATCCGTGAGTTTA	ACGACGTTATTGCGATGGAC			41	123	**25**

Table 6. Continued

Contig_Position	Primers	Allele Counts for 82 unique genotypes					
SNP IDs	Forward	Reverse	A	C	G	T	% Alternative Allele
133_1010	ATTATGGCGCTCTGCTCCTA	CCGTCTATCTGCCAAGAAGG	122		42		**26**
712_2077	AGTGCTGCTGGATCAGTGG	CTATCAGCGTCAGTGGCAAG	42		122		**26**
37_30149	ATACGCAGAAATGGGAGCAA	TCAGCATGTCCTCTTCCTCA	44		120		**27**
157_954	CTACGCAGTTAAGCGCTGTG	ATAGCGTGCTTAGCGGTGTT	119		45		**27**
301_957	CGTCAACGTACAGCTGCATT	TCGTCATTCGTTCCTCTTCA	47		117		**29**
150_21870	TGGAGATGCTGGTAACGATG	GTCTTCCCGTTGGTAGCATA	55		109		**34**
233_2053	CACTTGGAGCTCCTGGAAAC	GTGCTTAAACGGTGGTGGTC		58		106	**35**
1054_12307	ATCGCAGCTGCTCATTCTCT	TCTGCTGACTTCGAGAACCA		62		102	**38**
162_20472	ACGTTGCTTCGGTCTCAAGT	TCCAAGAGAGTGGAACTGGAA			101	63	**38**
863_51377	TGGTTTCCATCTCGTCATCA	AATCATGTATGGCTCGACGAC		65		99	**40**
638_43925	CTGTCACCAAGTGCACGACT	TTGCCACCTGAACGACTGTA	97		67		**41**
851_9386	AAGATGCTGTCCGGGAGATT	TACCTTCTCCACCACCGTGT			97	67	**41**
154_12147	CTCTCCATGACAAGGTGTGC	ATATCATGGCGGACACACG			69	95	**42**
405_1475	TTCGATAACTTGCACCAGGA	ATCAGCCGAGCTGAACTTGT		69		95	**42**
247_11808	TTCGTATGCTGCGCTAACAA	GCTTTTTCGCTGGCAATAAC	70	94			**43**
14_20241	AGCAAGAAATTTGGGGACAA	AGCTCGAATGTTCGGAAAAA		93		71	**43**
194_1933	CCACATACTCAACGGCTTCA	CCAGATGCCGAGCTAATGTT	72		92		**44**
516_687	CGCACGTTTTCGTCTTTTGT	AGCTGTGGCCAGAAGGATTA	76		88		**46**
505_1292	CAGACGACCATGTTGATTGG	TTTCAAAGTGGTCCACGACA		87		77	**47**
524_14500	CTGATGCCGCATACTGAGAC	TCAAATCCGTGCAACACAAT	87	77			**47**
425_3047	TCCACAGGTGCATGTTCTTC	TATGATGTGGAGCTCGATGG	86		78		**48**
6_18952	GCTTATCGGTGAACACAACG	TTATCTTTCGCCCACATGGT		84		80	**49**

Vita

Rebecca Lyon graduated cum laude with honors research distinction from the Ohio State University in 2012 with a Bachelor of Science degree in Crop Science. She completed her undergraduate thesis on the role of type III effectors in disrupting guard cell function that prevents bacterial invasion. In fall 2012, she was awarded the J. Wallace & Katie Dean Graduate Fellowship and joined the Department of Entomology and Plant Pathology at the University of Tennessee, Knoxville.